Static Non-Linearity Based Analysis of Non-Resonant MOSFET as THz Detector

Muhammad Ibrahim Wasiq Khan, Suna Kim, *Member, IEEE,* Dae-Woong Park, Hyoung-Jun Kim, Seok-Kyun Han, and Sang-Gug Lee, *Member, IEEE*

***Abstract*—The non-resonant THz response of CMOS FET has been analyzed based on static non-linearities of the transistor channel. Under the quasi-static limit, the second order non-linearities dominantly determine the DC current in the channel generated in response to THz signal. For the applied gate-to-source and drain-to-source signals, the significance of the second order non-linear terms is analyzed as a function of DC bias conditions. Based on the analysis, it has also been shown that a differential detector topology can give highest responsivity in cold operation. A tradeoff between the channel resistance and load impedance plays a vital role in detecting the generated current. Moreover, the design considerations such as readout modes, optimum bias points and device dimensions are also discussed with respect to both responsivity and noise equivalent power (NEP).**

***Index Terms*—THz detection, plasma wave detection, MOSFET power detector, transistor channel non-linearity, loading effects, differential detector configuration**

I. Introduction

TERAHERTZ frequency range $(0.3 - 3\ THz)$ has been an active research area for its numerous potential applications in security, biomedical imaging, spectroscopy and high-speed data transmission [1]-[3]. For all these applications, the THz detector is a fundamental building block. Several detectors based on Schottky diode, compound semiconductors, strained- and bulk-silicon CMOS have been presented [4]-[9]. CMOS based THz detectors has attracted a great deal of attention for their competitive performance at room temperature, low cost, high yield and capability of integration with other signal processing circuits.

The operating principle of transistors as THz detector can be explained by the plasma-wave [10] or distributive mixing [11] theories. The plasma-wave detection theory, introduced by Dyakonov-Shur, explains the response of HEMT to THz electromagnetic radiations [12]. It regards the channel of a transistor as a two dimensional (2D) electron gas that as a whole exhibits hydrodynamic behavior. Thus plasma waves generated in the channel by an external THz source can produce DC voltage or current in the channel. The distributed resistive mixing theory meanwhile regards the channel of a transistor as a non-quasi-static RC-ladder, such that each segment of the channel operates as a self-mixing power detector.

The authors are with the Department of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 305-701,Korea, (e-mail: ibrahimwasiq@kaist.ac.kr)

This work was supported by the Pioneer Research Center Program through the National Research Foundation of Korea funded by the Ministry of Science, ICT & Future Planning under Grant 2012-0009594.

Assuming the electron mobility $\mu = 0.0331\ m^2/Vs$ for typical silicon CMOS, the moment relaxation time $\tau = \mu m/|e|$ is 49 fs, where m is effective transport mass $(m = 0.259m_0)$ and e the electron charge [13]. Thus $\omega\tau < 1$ for terahertz gap $(\omega/2\pi = 0.3 - 3\ THz)$ and CMOS FET works as non-resonant detector in this regime. The quasi-static limit of transistor operation (where transistor can be modeled by lumped elements) is given by the condition that the plasma-wave transit time τ_t^{pl} through the channel is much shorter than the oscillation period of radiation field $(\tau_t^{pl} \ll 2\pi/\omega)$ [13]. For $e(V_{GS} - V_{th}) \gg \eta k_B T$, plasma-wave transit time is given by

$$\tau_t^{pl} = \sqrt{\frac{L^2 m}{e(V_{GS} - V_{th})}} \tag{1}$$

where η is the ideality factor of transport in the channel, k_B the Boltzmann constant, T the temperature, V_{th} the threshold voltage for transistor, V_{GS} the gate-source bias voltage and L the channel length. For 65 nm CMOS transistor, with $(V_{GS}-V_{th}) = 0.1\ V$, plasmon transit time evaluates to $0.25\ ps$ whereas the oscillation period for incident $500\ GHz$ signal is $2\ ps$. Thus, we are at the edge of quasistatic limit. In this paper, we will present that even at $500\ GHz$ the quasi-static analysis considering the non-linearity of transistor channel can closely predict the non-resonant THz response.

The analysis is presented in Section II. Three dimensional plots for voltage and current responsivity and noise equivalent power (NEP) are also included. Optimum detector topology and device dimensions are described in Section III. Section IV offers the implementation details of the detector. Section V compares the simulated results with the measurement ones. Section VI concludes.

II. Analysis of THz Power Detector

From the non-linear behavior of the MOS transistor, the small-signal drain-to-source current (i_{ds}) is a function of the gate-to-source (v_{gs}) and drain-to-source (v_{ds}) voltages and can be expressed by the two-dimensional Taylor series as

$$\begin{aligned} i_{ds}(v_{gs}, v_{ds}) =& g_{m1}v_{gs} + g_{ds1}v_{ds} + g_{m2}v_{gs}^2 + g_{ds2}v_{ds}^2 + \\ & g_{11}v_{gs}v_{ds} + g_{m3}v_{gs}^3 + g_{ds3}v_{ds}^3 + \\ & g_{12}v_{gs}v_{ds}^2 + g_{21}v_{gs}^2 v_{ds} + \cdots \end{aligned} \tag{2}$$

where the Taylor series coefficients are obtained from the derivatives of DC drain-to-source current I_{DS} as

$$g_{mx} = \frac{1}{x!}\frac{\partial^x I_{DS}}{\partial V_{GS}^x}\bigg|_{V_{DS}} \tag{3}$$

$$g_{dsx} = \frac{1}{x!}\frac{\partial^x I_{DS}}{\partial V_{DS}^x}\bigg|_{V_{GS}} \tag{4}$$

$$g_{xy} = \frac{1}{x!y!}\frac{\partial^{x+y} I_{DS}}{\partial V_{GS}^x \partial V_{DS}^y} \tag{5}$$

If $v_{gs} = V_{THz}\cos(\omega_{THz}t)$ and $v_{ds} = \alpha V_{THz}\cos(\omega_{THz}t + \theta)$ (where α represents the amplitude ratio of v_{ds} with respect to v_{gs} and θ the phase difference between them), it can be shown that the even order non-linearity generates a DC component. All other frequency components (at integer multiple of ω_{THz}), being located well beyond the f_{max} of the transistor, are heavily attenuated. Here, we will restrict the analysis to the second order terms as they are the dominant one and give square law power detection. Then, the DC current response to terahertz signal can be given by

$$I_{THz} = \frac{1}{2}(g_{m2} + \alpha^2 g_{ds2} + \alpha\cos\theta g_{11})V_{THz}^2 \tag{6}$$

We can also assume $v_{ds} = V_{THz}\cos(\omega_{THz}t)$ and $v_{gs} = \alpha V_{THz}\cos(\omega_{THz}t + \theta)$ without losing the generality. The coefficients described in equations (3),(4) and (5) can be obtained from the transistor model. In this work, BSIM4 is used to extract the Taylor series coefficients which were also verified from the measured IV curves. Fig. 1, 2 and 3 presents the behavior of second order non-linear terms: transconductance 'g_{m2}', output conductance 'g_{ds2}' and intermodulation 'g_{11}' respectively, as a function of DC bias conditions for $W/L = 2.4\mu m/60nm$ NMOS.

From Fig. 1, 2 and 3, it can be seen that there are two distinct regions of operation with respect to drain-source bias. One is unbiased drain-source or cold region where $V_{DS} \approx 0\ V$ and the other is drain-biased region, especially where $V_{DS} > 0.25\ V$. From now on when we refer to drain-biased operation, it means the region where $V_{DS} > 0.25\ V$. It is clear that in drain-biased case, g_{m2} is dominant whereas g_{ds2} and g_{11} are prominent in unbiased-drain operation ($V_{DS} \approx 0\ V$). Fig. 2 and 3 show that g_{ds2} and g_{11} have comparable magnitudes at $V_{ds} \approx 0$ but g_{ds2} is negative while g_{11} is positive. Thus, they generate current in opposite directions which are then subtracted. The peaks of g_{ds2} and g_{11} occurs at slightly different values of V_{GS}, therefore, the generated current can be maximized in between these peaks.

When drain-bias ($V_{DS} > 0.25\ V$) is applied, g_{m2} dominates the generated current. It can be seen from Fig. 1 that for $V_{GS} \leq V_{th}$, g_{m2} continues to increase with the increase in V_{DS} whereas for $V_{GS} > V_{th}$, it shows a peak at certain V_{DS} and then start to decrease. Hence, the generated current follows the same trend. It should be noted that the magnitude of g_{m2} is significantly smaller than that of g_{ds2} and g_{11}, but because of opposite polarity of g_{ds2} and g_{11} (assuming $\theta = 0°$), those currents tend to cancel each other in cold operation whereas the currents of g_{m2} and g_{11} tends to add in drain-biased region. Therefore, the generated current shows little increment going from cold region to drain-biased region.

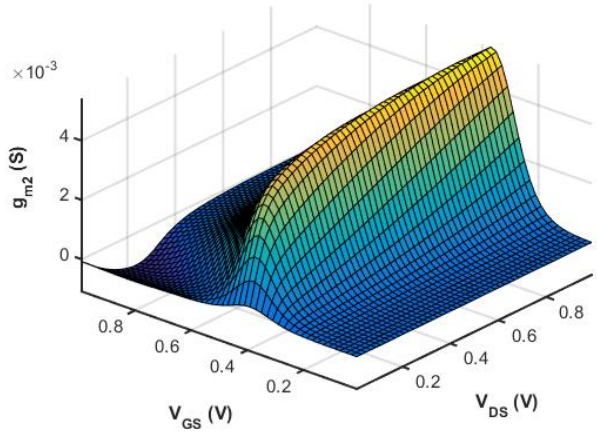

Fig. 1. g_{m2} versus V_{GS} and V_{DS}

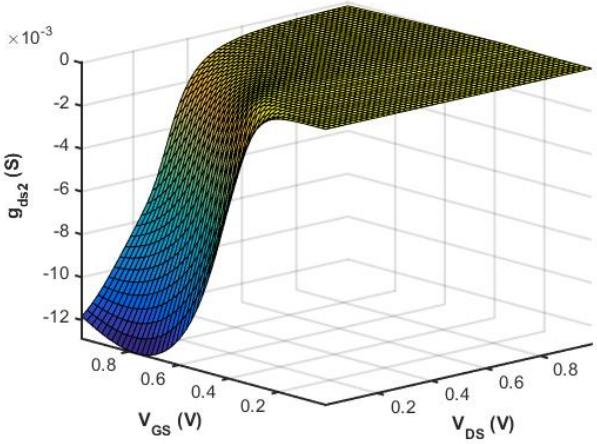

Fig. 2. g_{ds2} versus V_{GS} and V_{DS}

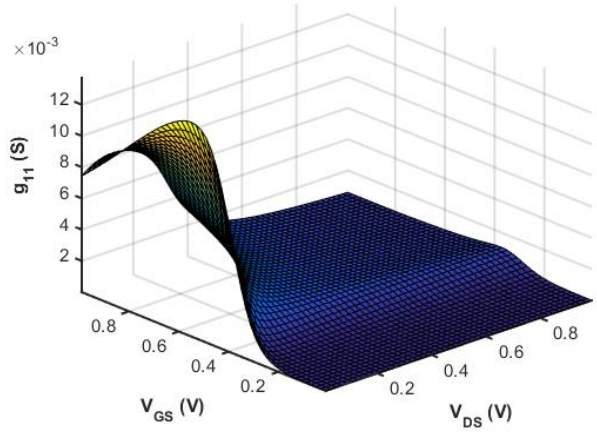

Fig. 3. g_{11} versus V_{GS} and V_{DS}

A. Current and Voltage Responsivity

The way the DC current, shown in (6), is extracted out of the detector depends on the read out circuit. Therefore, in order to complete the analysis, it is necessary to include its effect. The equivalent circuit of the THz detector including the load is shown in Fig. 4, where R_g is the gate resistance and R_{nqs} represents the non-quasi static effect of the transistor. C_{gs}, C_{gd} and C_{db} are the gate-to-source, gate-to-drain and drain-to-base parasitic capacitances, respectively. Assuming that the MOSFET is conjugate matched with the THz source, we can write $V_{THz}^2 = P_{in}R_{in}$, where P_{in} is the available input power and R_{in} the real part of the source impedance. For the load of $R_L||1/\jmath\omega_m C_L$, the current responsivity is given by

$$R_i = \frac{i_{out}}{P_{in}} = \frac{I_{THz}\left[\frac{1+\jmath\omega_m C_L R_L}{1+(g_{ds}+\jmath\omega_m C_L)R_L}\right]}{P_{in}}$$

$$R_i = \frac{1}{2}\left(g_{m2} + \alpha^2 g_{ds} + \alpha\cos\theta g_{11}\right) R_{in} \left[\frac{1+\jmath\omega_m C_L R_L}{1+\left(g_{ds}+\jmath\omega_m C_L\right)R_L}\right] \quad (7)$$

where ω_m is the modulation frequency of the amplitude modulated THz signal. Similarly, voltage responsivity is given by

$$R_v = \frac{v_{out}}{P_{in}} = \frac{1}{2}\left(g_{m2} + \alpha^2 g_{ds2} + \alpha\cos\theta g_{11}\right) R_{in} \left(R_L \parallel \frac{1}{\jmath\omega_m C_L} \parallel \frac{1}{g_{ds}}\right) \quad (8)$$

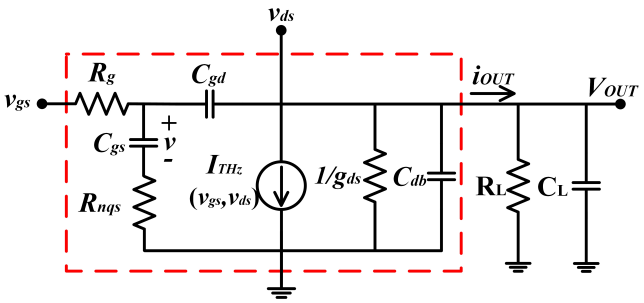

Fig. 4. Equivalent circuit of THz detector with load impedance

The calculated voltage and current responsivities, for all possible values of V_{GS} and V_{DS}, are plotted in Fig. 5 and 6 for the same $2.4\mu m/60nm$ NMOS FET (assuming $\theta = 0°$). One noticeable difference between the voltage and current responsivities is that the voltage responsivity is an order of magnitude higher in drain-biased region than in cold operation whereas the current responsivity shows little variation. This phenomenon can be explained by the voltage division between channel resistance and load impedance.

In case of current responsivity, since the load resistance is quite small compare to the channel resistance, the behavior of R_i is essentially determine by the current generated by the second order non-linearities only. The current responsivity is a little lower in unbiased-drain region than in drain-biased region because of the opposite polarity between g_{ds2} and g_{11}, as already explained earlier. In Fig. 6, the peak current responsivity occurs at $V_{GS} > V_{th}$ because the peaks of g_{m2}, g_{ds2} and g_{11} occur above V_{th}. The peak current responsivity also follows the same behavior as that of g_{m2} with respect to V_{DS} in drain-biased region.

For voltage responsivity in cold operation, the channel resistance is low and hence little voltage appears across load (Fig. 4). As drain voltage increases from zero, both the channel resistance and the generated current increases. Hence, the output voltage increases. It is important to note that the peak voltage responsivity occurs at $V_{GS} < V_{th}$ and moves towards lower V_{GS} with increase in V_{DS}. This behavior is again because of the voltage division between the channel resistance and load.

From the above analysis, one can easily find the optimum bias point for the maximum responsivity.

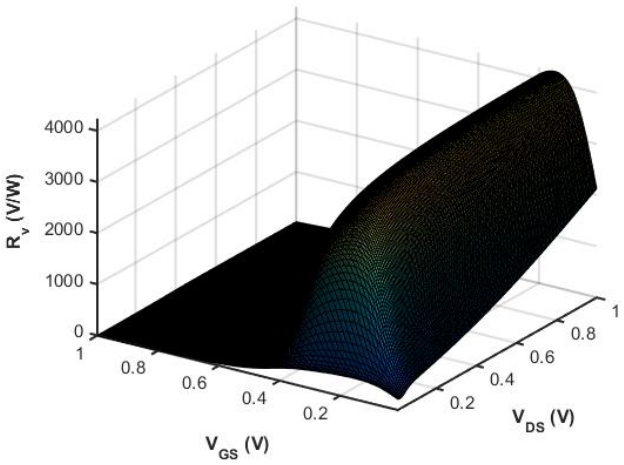

Fig. 5. Voltage responsivity for $2.4\mu m/60nm$ NMOS with $R_L = 10\ M\Omega$

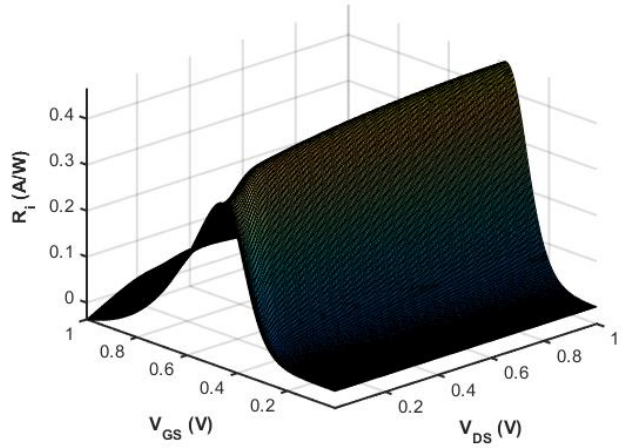

Fig. 6. Current responsivity for $2.4\mu m/60nm$ NMOS with $R_L = 50\ \Omega$

B. Noise Equivalent Power

The noise-equivalent power (NEP) is defined as the power at which the signal-to-noise ratio becomes unity for detector time constant of 1 sec. In this work, only the channel thermal noise is considered (even in the case where the MOSFET operates in saturation mode) for simplicity. The noise spectral power density of channel thermal noise is given by $\bar{i_n^2} = 4k_BT\gamma g_{ds0}$, where the coefficient $\gamma = 2/3$ and g_{ds0} is channel conductance at $V_{DS} = 0\ V$. The NEP of the detector only, in both the current and voltage read out mode, is essentially the same given by the ratio of channel current noise to the generated current response and can be expressed as

$$NEP = \frac{\sqrt{\bar{i_n^2}}\left[\frac{1+\jmath\omega_m C_L R_L}{1+(g_{ds}+\jmath\omega_m C_L)R_L}\right]}{R_i}$$

$$NEP = \frac{2\sqrt{4k_BT\gamma g_{ds0}}}{(g_{m2} + \alpha^2 g_{ds2} + \alpha\cos\theta g_{11})\, R_{in}} \tag{9}$$

The simulated NEP for $2.4\mu m/60nm$ NMOS is shown in the Fig. 9. The minimum NEP of 6.18 $pW/\sqrt{Hz}$ is obtained at $V_{DS} = 1\ V$ and $V_{GS} = 300\ mV$ ($< V_{TH}$) which is close to the peak point of voltage responsivity. The minimum NEP in cold operation is higher by more than two times, i.e. 14.7 $pW/\sqrt{Hz}$, at $V_{GS} = 420\ mV$ ($> V_{TH}$). This point is closer to the peak current responsivity. Thus, it can be said that if detector is in cold operation, then the current readout mode should be preferred as it can achieve low NEP and high current responsivity (Fig. 6) simultaneously for the same gate bias. Similarly, the voltage readout mode can provide minimum NEP and peak voltage responsivity (Fig. 5) simultaneously for the same gate bias, when detector is drain-biased.

It is important to note that with DC drain current, $1/f$ noise is also added which can severely degrade NEP. Therefore, the drain-biased detector should be used only when THz signal can be modulated above the $1/f$ noise corner frequency and the detector output can be high pass filtered.

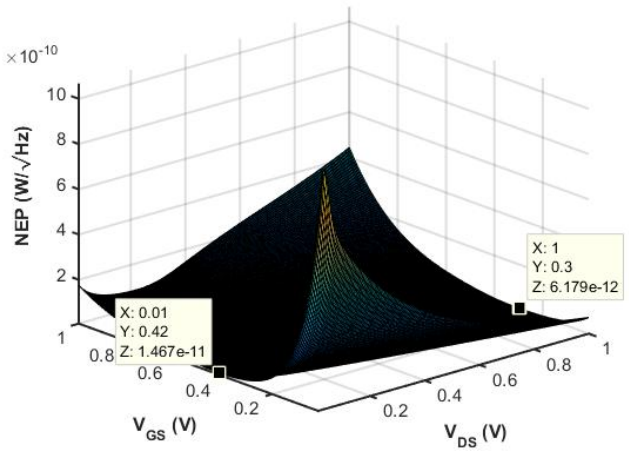

Fig. 7. Calculated NEP for $2.4\mu m/60nm$ NMOS

III. Responsivity Dependence on Detector Topology, Biasing and Dimensions

A. Detector Topology and Biasing

In this section, we will compare the response of four possible configurations of the THz detector with respect to the THz signal application, as shown in Fig. 8, where the gate-source and drain-source ac voltage are given along with respective circuit schematic. Here we will assume $\alpha = 1$ for all cases, and $\cos\theta = 1$ or -1 represents the in-phase and differential application of v_{gs} and v_{ds}, respectively.

In detector topology (a), the THz signal is applied at the gate side only, so the response will depend on g_{m2} which is very low in unbiased-drain region but high in drain-biased mode (Fig. 1).

Topology (b) is the most commonly used architecture. In this case, THz signal is coupled to both gate and drain through a large capacitor C_e. Therefore, all second order terms become important (equation (6)). The generated current and the effects of load are already explained in Section II. The voltage and current responsivities are shown in Fig. 5 and 6, respectively.

In topology (c), THz signal is applied to the drain node only and so transconductance non-linearity g_{ds2} matters. In cold operation, this topology generates higher current than topologies (a) and (b), because there is no opposite current generated from g_{11}. In drain-biased case, this topology will show the lowest responsivity of all four configurations because $g_{ds2} \approx 0$.

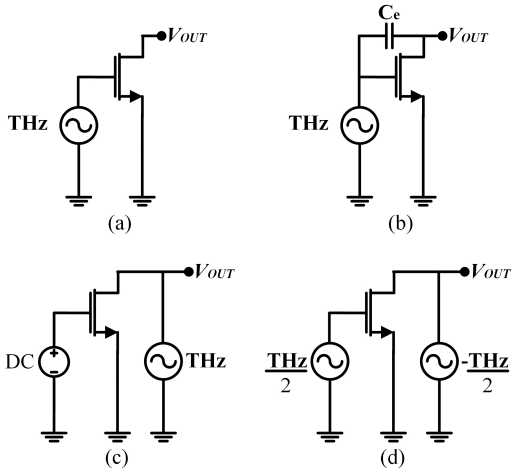

Fig. 8. Four possible topologies for THz detector (a) $v_{gs} = V_{THz}\cos(\omega_{THz}t)$, $v_{ds} \approx 0$, (b) $v_{gs} = v_{ds} = V_{THz}\cos(\omega_{THz}t)$, (c) $v_{gs} \approx 0$, $v_{ds} = V_{THz}\cos(\omega_{THz}t)$, (d) $v_{gs} = -v_{ds} = V_{THz}\cos(\omega_{THz}t)$

With topology (d) in cold operation, the differential signals at the gate and drain node is beneficial as g_{ds2} and g_{11} will add up instead of being cancelled. Thus, the differential derive of gate and drain nodes would lead the highest responsivity among all four topologies. This will also improve the NEP beyond what is shown in Fig. 7, since channel thermal noise remains the same. But in drain-biased case, the responsivity would be low because the currents generated by g_{m2} and g_{11} have opposite polarity. The actual implementation of a

differential topology poses several challenges (detector design is discussed in next section). If two MOSFETs are used with a differential antenna as detector front end, then two separate input matching networks are required for gate and drain sides. These matching networks can add different losses and delays to the THz signal, thus complicating the design.

The order of responsivities for unbiased-drain case is $(a) < (b) < (c) < (d)$ and for drain-biased case is $(c) < (d) < (a) < (b)$ as shown in Fig. 9 and Fig. 10, respectively. Thus, it can be concluded that in unbiased-drain and drain-biased regions, differential and in-phase signals should be used, respectively, for maximum responsivity.

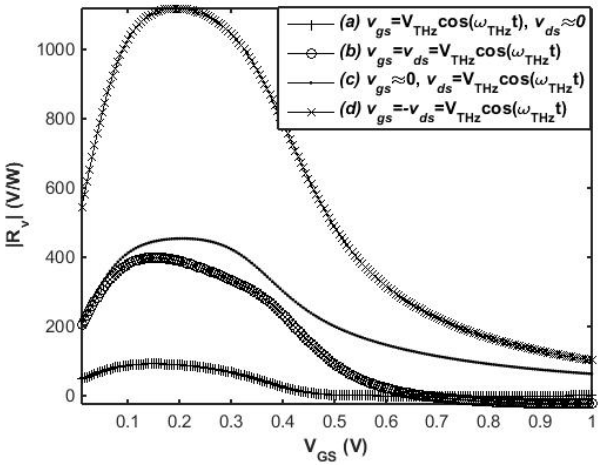

Fig. 9. Comparison of voltage responsivity for different topologies for $2.4\mu m/60nm$ NMOS at $V_{DS} \approx 0\ V$ and $R_L = 10\ M\Omega$

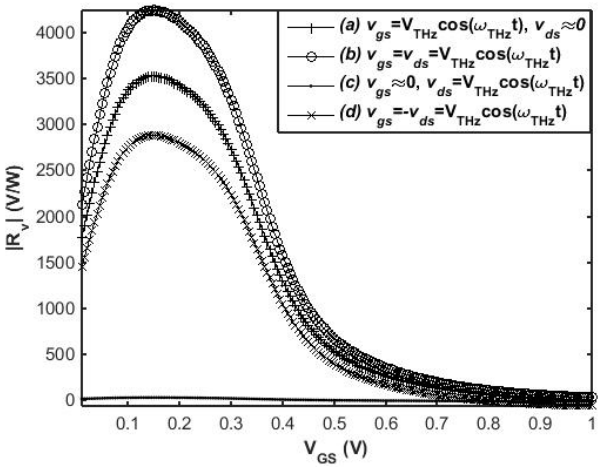

Fig. 10. Comparison of voltage responsivity for different topologies for $2.4\mu m/60nm$ NMOS at $V_{DS} = 1V$ and $R_L = 10\ M\Omega$

B. Detector Dimension

All the terms g_{m2}, g_{ds2} and g_{11} are directly proportional to the aspect ratio (W/L), so the wider transistors with minimum allowed length should maximize i_{THz}. The adoption of minimum channel length is also in accordance with the resonant regime $(\omega\tau > 1)$ of plasmon interference where the channel length approaches the critical decay length of plasma waves [12] thus increasing responsivity. However, in non-resonant regime, MOSFET dimension should be chosen considering the loading effects and channel thermal noise. To understand this, let's consider the equivalent circuit current and voltage readout modes as shown in Fig. 4. Detector with in-phase gate and drain signals is considered in the following analysis.

The peak values of calculated voltage responsivity at $V_{DS} = 1\ V$ for different dimensions are shown in Fig. 11. It is important to mention that this result is for $R_L = 10\ M\Omega$ (which is close to the ideal case of infinite load). Output voltage is dependent on both the generated current and the channel resistance. Higher W/L increases i_{THz} but decreases channel resistance at a higher rate. Thus, longer channels with narrower widths show better voltage responsivity. This has been verified up to 300 GHz through on-wafer measurements [14].

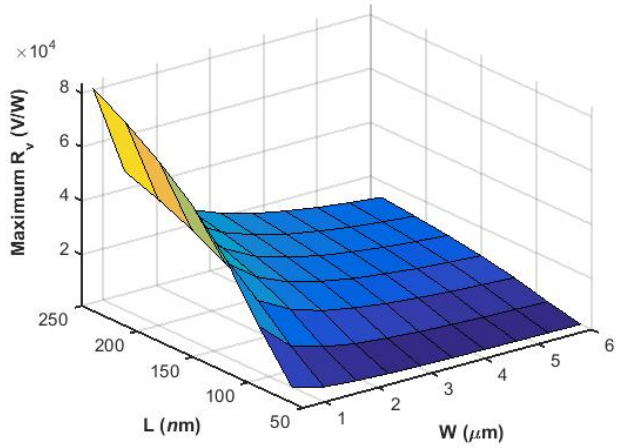

Fig. 11. Maximum calculated voltage responsivity for different lengths and widths of the MOSFET

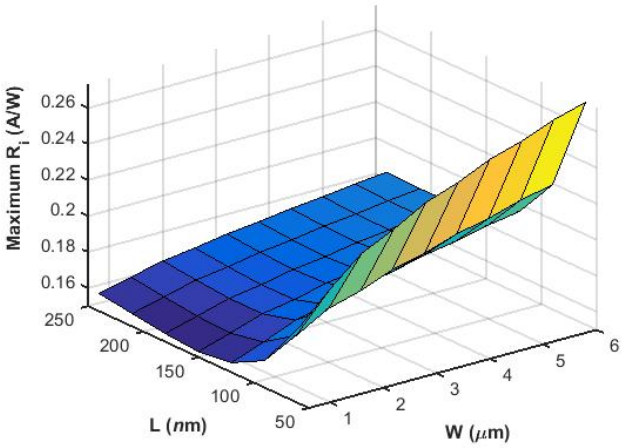

Fig. 12. Maximum calculated current responsivity for different lengths and widths of the MOSFET

The variation of current responsivity with device dimensions is as shown in Fig. 12 ($R_L = 50\ \Omega$, close to the ideal case of zero ohm load). In current readout mode, the improvement in responsivity with the aspect ratio is because of the enhancement in i_{THz}, , as long as $R_L \ll \frac{1}{g_{ds}}$. Thus, shorter channels with wider widths show better current responsivity.

The voltage and current readout modes show opposite tendency with W/L because of the variation of R_L from zero to infinity. If an intermediate value is chosen for R_L, then the behavior of voltage and current responsivities tends to become similar. Note that we haven't taken into account the effects of parasitic capacitances which may degrade the performance of detector (especially for high aspect ratios) by shunting the THz signal. However, by adopting the matching circuits, the effects of parasitic capacitances can be nullified, leading to similar responsivity behaviors predicted by this work.

As described in Section II, NEP is independent of loading effects (provided that the load adds negligible noise) and hence is a better figure of merit for choosing the device dimension. The minimum calculated NEP at $V_{DS} = 1\ V$ is shown in Fig. 13. Shorter channel lengths and narrower width detectors show better noise performance [14]. NEP also shows a peak with channel length variation for a fixed width. Increasing channel length from $60\ nm$ to $120\ nm$, increases NEP because of the more rapid reduction in the generated current compare to the channel thermal noise. From $120\ nm$ to $240\ nm$, the generated current decrease at a slower rate whereas the channel current noise reduces more sharply, hence, NEP improves. Reducing the width of the transistors, decreases the current noise at much faster rate than the reduction in generated current and thus narrower width leads to lower NEP.

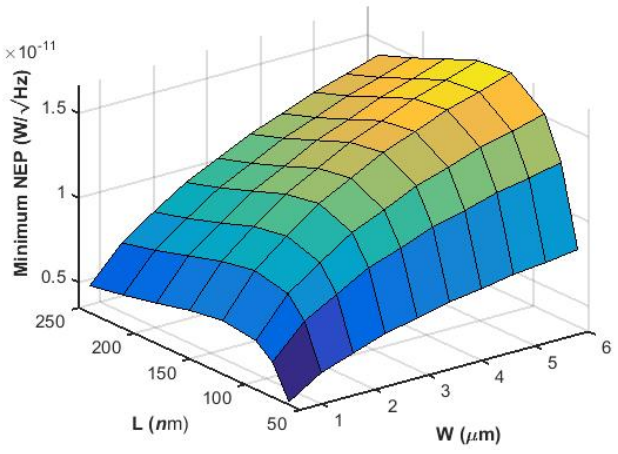

Fig. 13. Minimum calculated NEP for different lengths and widths of the MOSFET

IV. Detector Design

For verifying the analysis, MOSFET detectors are implemented in $65\ nm$ CMOS. Fig. 14 shows the circuit schematic and cross-sectional view of the detector along with the load impedance. As can be seen in Fig. 14, the detector adopts a differential patch antenna implemented in the top Aluminum layer and uses M6-M7 as the ground. Antenna directly feeds the two NMOS transistors (M1 and M2), connected together at their gates and drains respectively. The symmetry of the detector establishes a virtual ac ground at the V_{out} node and the output signal is extracted from this node. The quarter wavelength transmission lines (TL2) provides ac open at the drain terminals of MOSFETs while the open stub lines (TL1) are adopted for the impedance matching between the antenna and MOSFETs. The transmission lines are implemented as $1.5\ \mu m$ wide $25\ \Omega$ striplines in metal 4 that uses M1-M2 and M6-M7 as the grounds. The capacitor C_e is used to provide ac short between gate and drain. The gate DC bias is applied through the virtual ac ground of the antenna. As can be seen in Fig. 14(b), the whole circuit is confined under the antenna which allows area efficient implementation. Several detectors are fabricated with different transistor sizes and with or without the drain resistance R_D.

If the terahertz signal received by the antenna induces gate-to-source voltage $v_{gs} = V_{THz}\cos(\omega_{THz}t)$, then the drain-to-source voltage v_{ds} can be expressed as

$$v_{ds} = \left[\frac{\frac{1}{g_{ds}} \parallel \frac{1}{j\omega_{THz}C_{db}}}{\frac{1}{j\omega_{THz}C_e} + \left(\frac{1}{g_{ds}} \parallel \frac{1}{j\omega_{THz}C_{db}}\right)}\right] = \alpha v_{gs}\angle\theta \quad (10)$$

where

$$\alpha = \frac{\omega_{THz}C_e}{\sqrt{g_{ds}^2 + \{\omega_{THz}(C_e + C_{db})\}^2}}, \qquad \theta = \frac{\pi}{2} - \tan^{-1}\left[\frac{\omega_{THz}(C_e + C_{db})}{g_{ds}}\right] \quad (11)$$

Note that in Fig. 14, the NMOS pair is in parallel, so their generated currents are added whereas the channel resistance is halved. Moreover, R_{in} represents the real part of input impedance of antenna. These considerations must be taken into account when using equations (7) and (8).

V. Comparison of Measurement and Simulation Results

Measurement Setup

The measurement setup is shown in Fig. 15. The detector is exposed to the $5\ mW$ $500\ GHz$ THz source. The source is amplitude modulated (electrically chopped) by the $1\ kHz$ square-wave signal provided by the lock-in amplifier which also measures the detector response. The input impedance of lock-in amplifier is set to $R_L = 10\ M\Omega$ and $1\ k\Omega$ for voltage and current readout modes, respectively and $C_L = 25\ pF$.

The incident power on the detector can be determined from the knowledge of transmitted power (P_t), the transmitting antenna gain (G_t) and the effective area (A_eff) of the patch antenna using the Friis transmission equation

$$P_{in} = \frac{P_t G_t}{4\pi r^2} A_{eff} \quad (12)$$

where $A_{eff} = \left(D\lambda^2\right)/4\pi$, D being the directivity and λ the free-space wavelength. Using the simulated directivity of

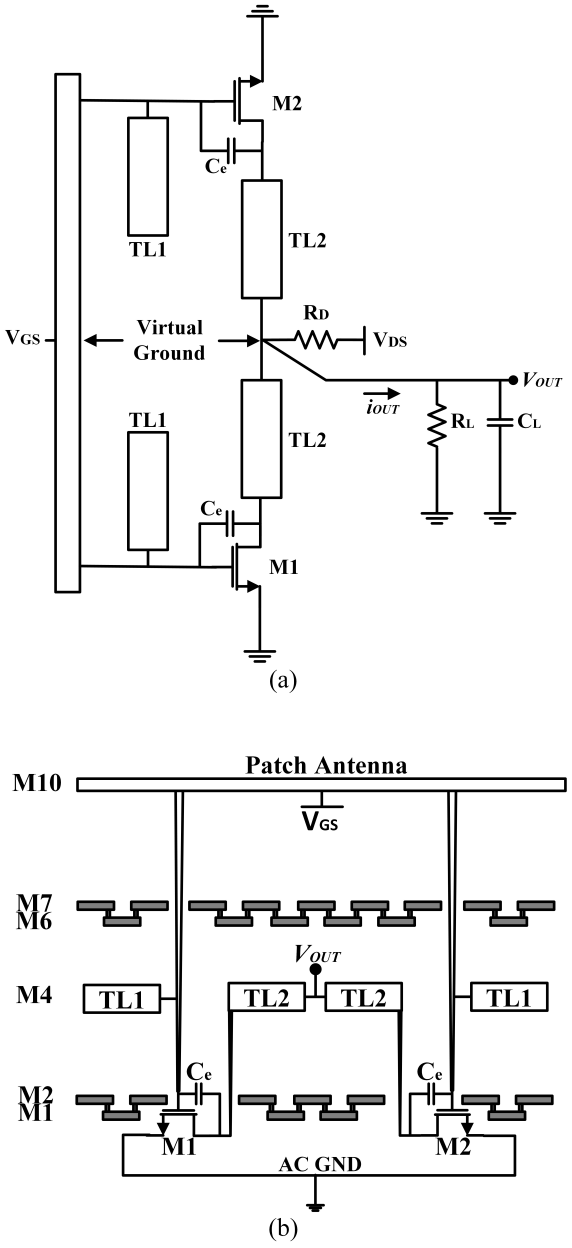

Fig. 14. (a) Schematic and (b) Cross sectional view of the THz detector

$6.7\ dB$, the estimated incident power at the antenna of the detector is $250\ nW$ at $1.5\ cm$ distance. With the simulated antenna radiation efficiency of 60%, the power delivered to the MOSFETs is $150\ nW$.

It is apparent from equations (7) and (8) that the comparison in either voltage or current readout mode is enough. Hence, we will present the measurement result of voltage response only.

A. Unbiased Drain (Zero DC Drain Current) Detector

In this case, g_{ds2} and g_{11} are involved. If $\omega_{THz}\,(C_e + C_{db}) \gg g_{ds}$ in equation (11), $\cos\theta \approx 1$ and $\alpha < 1$, the contribution of g_{ds2} becomes less compare to g_{11} (equation 6) and hence the generated current remain positive. On the other hand, if $g_{ds} \ll \omega_{THz}\,(C_e + C_{db})$,

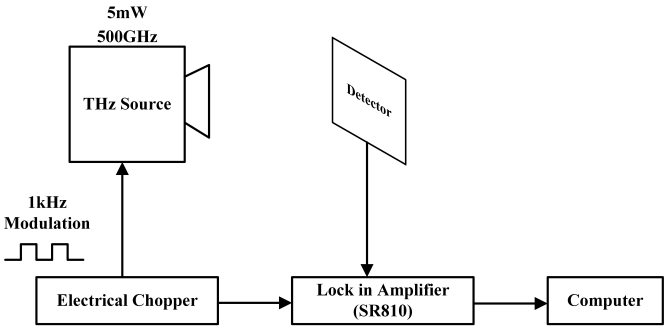

Fig. 15. Measurement setup for responsivity

$\cos\theta \approx 0$ and $\alpha \ll 1$, there is no contribution of g_{11} while that of g_{ds2} is heavily attenuated. Thus, an ac short between gate and drain through a large capacitor C_e is a necessary condition for proper operation only when the detector is in cold operation and THz signal is applied at the gate terminal.

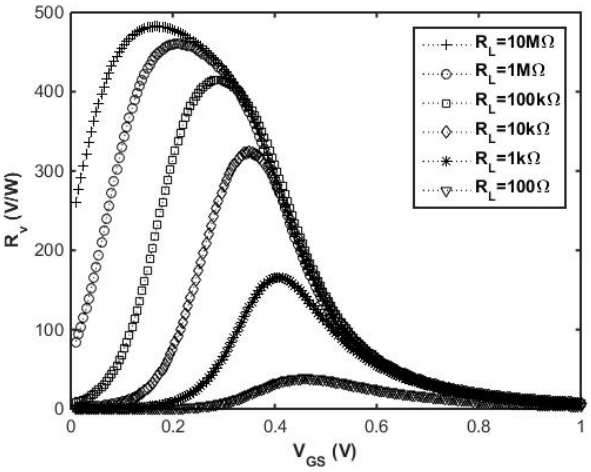

Fig. 16. Calculated voltage responsivity for $2.4\mu m/60nm$ NMOS at $V_{DS} = 0\ V$ and varying loads

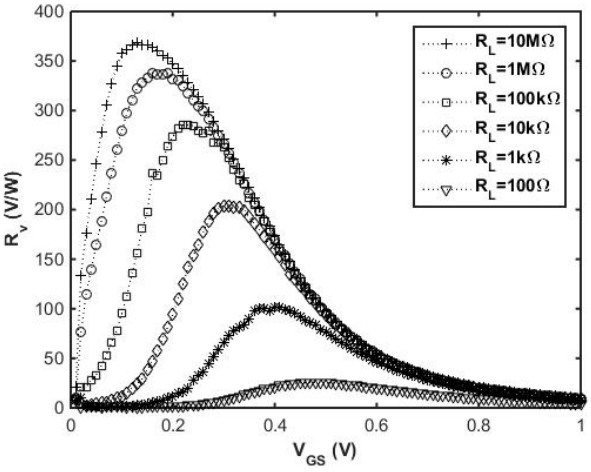

Fig. 17. Measured voltage responsivity for $2.4\mu m/60nm$ NMOS at $V_{DS} = 0\ V$ and varying loads

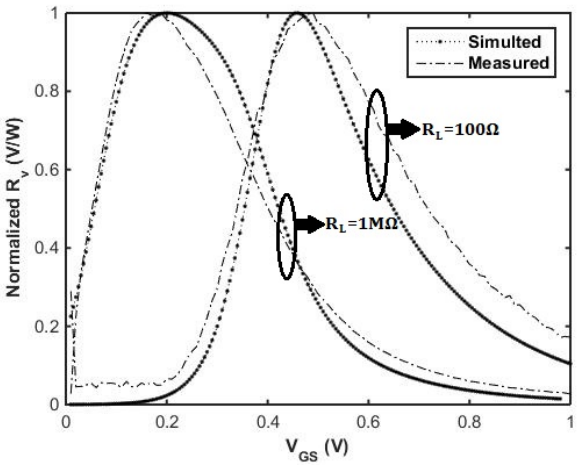

Fig. 18. Measured and calculated normalized voltage responsivity for $2.4\mu m/60nm$ NMOS at $V_{DS} = 0\ V$ and $R_L = 1\ M\Omega$ and 100 Ω

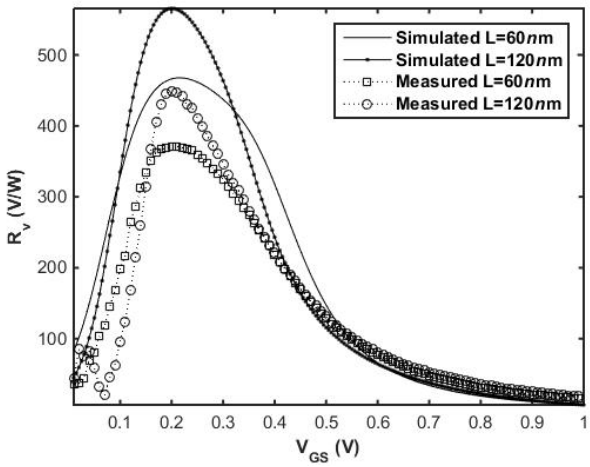

Fig. 19. Dependence of voltage responsivity on length of the MOSFET $(W = 0.6\mu m)$

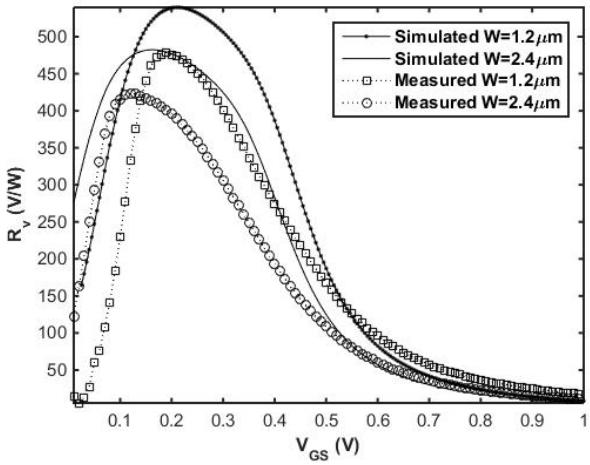

Fig. 20. Dependence of voltage responsivity on width of the MOSFET $(L = 60nm)$

In order to verify the above analysis, the MOSFETs in the detector are biased with open-drain while varying the load resistance from 100 Ω to $10M\ \Omega$. The calculated and measured voltage responsivities are shown in Fig. 16 and 17. The responsivities calculated using equation (8) well reflect the loading effects and it clearly shows that decreasing the load impedance will decrease the voltage responsivity. Although the generated current is determine by g_{ds2} and g_{11}, the peaking behavior of the output voltage is influenced by the parallel combination of channel resistance and load impedance. The peak responsivity occurs where two impedances are comparable. Above the peak bias point($V_{GS} > V_{GS-peak}$), channel resistance functions as internal shunt. Moreover, α and $\cos\theta$ also decrease above the $V_{GS-peak}$ and so does the generated current. Below the peak bias point, the generated current itself decreases because both g_{ds2} and g_{11} decreases. . Two of the normalized responsivities for $R_L = 100\ \Omega$ and $1\ M\Omega$ are also presented in Fig. 18 to show that equation (8) quite closely predicts the optimum gate-to-source voltage for peak responsivity.

The dependence of voltage responsivity over the channel length and width variation is shown in Fig. 19 and 20, respectively. It can be seen that, although the measured values are lower than the predicted ones, the behavior in the variation of responsivity with L and W agrees well. The reason for this behavior is already discussed in Section III. Thus, the proposed model can be used in predicting the device dimension for the optimum detector performance.

B. Drain-Biased (Current Biased) Detector

The drain bias is applied through an external $R_D = 10\ k\Omega$ resistor (Fig. 14) while keeping the gate bias constant. The comparison of the measured and calculated results is presented in Fig. 21. In this case R_D appears in parallel to R_L and hence effective load resistance is $\approx\ 10\ k\Omega$. For $V_{GS} < V_{th}$, the voltage response is determined by the parallel combination of the channel and load resistances multiplied by the generated current. Therefore, the voltage response will continue to increase with V_{DS} (not shown in Fig. 21). For $V_{GS} > V_{th}$, the channel resistance is small compare to that of load. Therefore, as can be seen from Fig. 21, the responsivity at $V_{GS} = 450\ mV$ is higher than that of $V_{GS} = 550\ mV$. The peak point is still dominantly determined by g_{m2}. The overall behavior of the model agrees well with the measurements.

C. Diode-Connected Detector

The detector with diode-connected NMOS is also fabricated, where the capacitor C_e is replaced with short between gate and drain and DC bias is applied through R_D. In this case, $v_{gs} = v_{ds}$ making α and $\cos\theta$ equal to 1 and equation (8) reduces to

$$R_v = g_{dio} R_{in} \left(R_D \parallel R_L \parallel \frac{1}{2g_{ds}} \right) \tag{13}$$

where $g_{dio} = \left.\frac{\partial^2 I_{DS}}{\partial V^2}\right|_{V_{GS}=V_{DS}}$. The comparison of the measured and calculated results is presented in the Fig. 22. The responsivity could only be measured up to $0.45\ V$, because

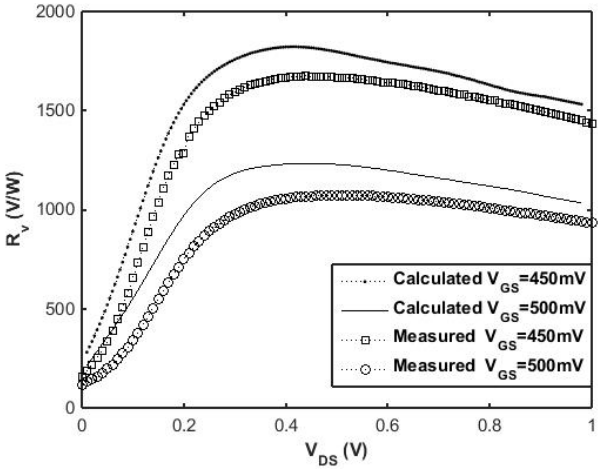

Fig. 21. Measured and calculated voltage responsivity for $2.4\mu m/60nm$ NMOS at $V_{GS} = 450mV$ and $500mV$

above this value the channel resistance shunt the additional DC voltage applied through R_D. A diode-connected device ensures the same voltage swing at gate and drain and provides easy biasing scheme but it loses the freedom of separately choosing optimum gate and drain bias points.

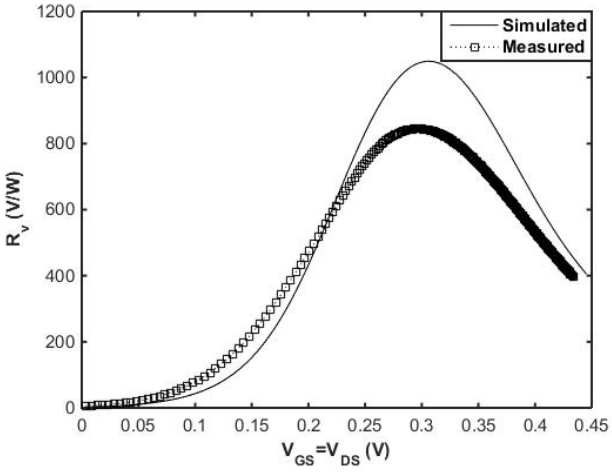

Fig. 22. Measured and calculated voltage responsivity of diode-connected NMOS detector $(2.4\mu m/60nm)$

Note that, overall, the measured responsivities are lower than what analysis predicts. This discrepancy can be explained by the fact that the proposed model did not account for the loss caused by the gate resistance and non-quasi static effects (represented by a resistance R_{nqs} in Fig. 4). The voltage signal v across C_{gs} is responsible for generating the current instead of v_{gs} as shown in Fig. 4. A more accurate analysis can be done by including dynamic non-linearity and non-quasi static effects. The chip micrograph showing six different patterns for THz detector is shown in Fig. 23.

VI. Conclusion

This work reports the non-resonant response of CMOS THz detectors based on the analysis of the static non-linearities of

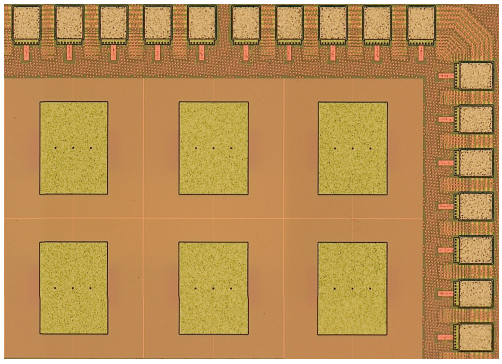

Fig. 23. Chip Micrograph $(980\mu m \times 700\mu m)$

the transistor channel. The second-order non-linearities of the drain current determine the DC current generated in response to THz signal. Loading effects are also considered in the analysis verified by comparison with measurement results. It has been shown that the combination of channel resistance and load impedance causes the voltage responsivity to become lower in unbiased-drain operation as compared the drain-biased, whereas the current responsivity shows little variations with drain-source bias as it is fairly independent of channel resistance. It has been predicted that driving a MOSFET detector with differential signals can maximize the responsivity in unbiased-drain operation because of the addition of the two second-order terms (g_{ds2} and g_{11}). Moreover, the model has been used to find out the optimum device dimensions for low NEP and high responsivity.

References

[1] M. Tonouchi, "Cutting-edge terahertz technology," Nat. Photonics, vol. 1, pp. 97–105, 2007

[2] P. H. Siegel, "Terahertz technology," IEEE Trans. Microw. Theory Tech., vol. 50, no. 3, pp. 910–928, 2002.

[3] S. E. Clark, J. A. Lovberg, C. A. Martin, V. Kolinko, and T. E. Corporation, "Passive millimeter-wave imaging for airborne and security applications," vol. 5077, pp. 16–21, 2003.

[4] R. Han, Y. Zhang, D. Coquillat, H. Videlier, W. Knap, E. Brown, and K. O. Kenneth, "A 280-GHz Schottky diode detector in 130-nm digital CMOS," IEEE J. Solid-State Circuits, vol. 46, no. 11, pp. 2602–2612, 2011.

[5] M. S. Shur and Jian-Qiang, "Terahertz sources and detectors using two-dimensional electronic fluid in high electron-mobility transistors," IEEE Trans. Microw. Theory Tech., vol. 48, no. 4 PART 2, pp. 750–756, 2000.

[6] S. Boubanga-Tombet, F. Teppe, D. Coquillat, S. Nadar, N. Dyakonova, H. Videlier, W. Knap, A. Shchepetov, C. Gardes, Y. Roelens, S. Bollaert, D. Seliuta, R. Vadoklis, and G. Valusis, "Current driven resonant plasma wave detection of terahertz radiation: Toward the Dyakonov-Shur instability," Appl. Phys. Lett., vol. 92, no. 21, pp. 1–4, 2008.

[7] S. L. Rumyantsev, K. Fobelets, D. Veksler, T. Hackbarth, and M. S. Shur, "Strained-Si modulation doped field effect transistors as detectors of terahertz and sub-terahertz radiation," Semicond. Sci. Technol., vol. 23, no. 10, p. 105001, 2008.

[8] W. Knap, F. Teppe, Y. Meziani, N. Dyakonova, J. Lusakowski, F. Boeuf, T. Skotnicki, D. Maude, S. Rumyantsev, and M. S. Shur, "Plasma wave detection of sub-terahertz and terahertz radiation by silicon field-effect transistors," Appl. Phys. Lett., vol. 85, no. 4, pp. 675–677, 2004.

[9] R. Tauk, F. Teppe, S. Boubanga, D. Coquillat, W. Knap, Y. M. Meziani, C. Gallon, F. Boeuf, T. Skotnicki, C. Fenouillet-Beranger, D. K. Maude, S. Rumyantsev, and M. S. Shur, "Plasma wave detection of terahertz radiation by silicon field effects transistors: Responsivity and noise equivalent power," Appl. Phys. Lett., vol. 89, no. 25, pp. 2004–2007, 2006.

[10] M. Dyakonov and M. Shur, "Shallow water analogy for a ballistic field effect transistor: New mechanism of plasma wave generation by dc current," Phys. Rev. Lett., vol. 71, no. 15, pp. 2465–2468, 1993.
[11] E. Ojefors, U. R. Pfeiffer, A. Lisauskas, and H. G. Roskos, "A 0.65 THz focal-plane array in a quarter-micron CMOS process technology," IEEE J. Solid-State Circuits, vol. 44, no. 7, pp. 1968–1976, 2009.
[12] M. Dyakonov and M. Shur, "Detection, mixing, and frequency multiplication of terahertz radiation by two-dimensional electronic fluid," IEEE Trans. Electron Devices, vol. 43, no. 3, pp. 380–387, 1996.
[13] A. Lisauskas, U. Pfeiffer, E. Öjefors, P. H. Bolvar, D. Glaab, and H. G. Roskos, "Rational design of high-responsivity detectors of terahertz radiation based on distributed self-mixing in silicon field-effect transistors," J. Appl. Phys., vol. 105, no. 11, 2009.
[14] S. Kim, D. W. Park, K. Y. Choi, and S. G. Lee, "MOSFET Characteristics for Terahertz Detector Application from On-Wafer Measurement," IEEE Trans. Terahertz Sci. Technol., vol. 5, no. 6, pp. 1068–1077, 2015.

Muhammad Ibrahim Wasiq Khan received the B.S. degree in Electrical engineering from National University of Science and Technology (NUST), Islamabad, Pakistan, in 2012, and the M.S. degree in Electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2016. His research interests include mm-Wave and THz integrated circuits especially THz detectors and THz imaging systems based on CMOS technology.

Suna Kim (M'15) received the B.S, M.S. and Ph.D. degrees in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2006, 2009 and 2016, respectively. She is now with Samsung Electronics. Her research interests include RF integrated circuit designs and further digital calibration circuit designs for wireless transceivers. Her research interests now extend to terahertz (THz) devices and circuit designs especially for THz detectors and THz imaging systems based on CMOS technology.

Dae-Woong Park was born in Daegu, Korea. He received the B.S. degree from Sungkyunkwan University, Seoul, Korea, in 2013, and the M.S. degree from Korea Advanced Institute of Science and Technology (KAIST), Daejon, Korea, in 2015, where he is currently working toward the Ph.D. degree at the Department of Electrical Engineering. His interests include device physics and extremely high-frequency (THz) CMOS devices and circuits.

Hyoung Jun Kim was born in Anyang, Korea in 1991. He received the B.S, degree in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2015, where he is currently working the M.S. degree. His research interests include design and analysis of Terahertz frequency multiplier.

Seok-Kyun Han received the B.S. degree in electronics engineering from Kwangju University, Korea, in 1995, and the M.S. and Ph.D. degrees in electrical engineering from Mokpo University and Mokpo Maritime University, Korea, in 1998 and 2004, respectively. From 1995 to 1998, he was with Information and Communications University, Daejeon, Korea, as a research assistant professor. Since 2010, he has been with the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in the Department of Electrical Engineering as a Research Associate Professor. His research interests are the CMOS-based RFIC and Microwave circuit designs.

Sang-Gug Lee (M'09) received the B.S. degree in Electronic engineering from Kyungpook National University, Daegu, Korea, in 1981, and the M.S. and Ph.D. degrees in Electrical engineering from the University of Florida, Gainesville, FL, USA, in 1989 and 1992, respectively. In 1992, he joined Harris Semiconductor, Melbourne, FL, USA, where he was engaged in silicon-based RFIC designs. From 1995 to 1998, he was with Handong University, Pohang, Korea, as an Assistant Professor with the School of Computer and Electrical Engineering. From 1998 to 2009, he was with the Information and Communications University, Daejeon, Korea, and became a full Professor. Since 2009, he has been with the Department of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, as a Professor. His research interests include CMOS-based RF, analog, and mixed-mode IC designs for various radio transceivers, especially the ultra-low power applications. Lately, his research interests extend to extreme high-frequency (THz) circuit designs, display semiconductors, and energy- harvesting IC designs.